Adding Global Soils Data to the Automated Geospatial Watershed Assessment Tool (AGWA)

L.R.Levick[1], D.J.Semmens[2], D.P.Guertin[3], I.S.Burns[1], S.N.Scott[1], C.L.Unkrich[1], D.C.Goodrich[1]

Abstract
The Automated Geospatial Watershed Assessment Tool (AGWA) is a GIS-based hydrologic modeling tool that is available as an extension for ArcView 3.x from the USDA-ARS Southwest Watershed Research Center (www.tucson.ars.ag.gov/agwa). AGWA is designed to facilitate the assessment of land-use and climate-change impacts on water yield and quality at multiple scales. It parameterizes two watershed runoff and erosion models, the Kinematic Runoff and Erosion Model (KINEROS2) and the Soil and Water Assessment Tool (SWAT), using readily available topographic, soils, and land-cover data. After parameterization, the selected model is run through the interface, and results are imported back into the GIS for display and analysis. AGWA was originally designed to obtain hydrologic parameters from the State Soil Geographic (STATSGO) and Soil Survey Geographic (SSURGO) databases, which are only available for the United States. The latest version of AGWA (1.4x) has incorporated the ability to derive inputs from the internationally available Food and Agriculture Organization of the United Nations (FAO) digital soil map of the world. The ability to use FAO soils in AGWA facilitates the analysis of trans-boundary watersheds by avoiding difficulties associated with different classification schemes on either side of the border. When used with existing global, classified land-cover maps it is now easy to run hydrologic simulations for trans-border watersheds. The structure and organization of the FAO soils dataset is fundamentally different from the STATSGO and SSURGO datasets since it covers the entire globe and must describe a wider range of soils with a more generalized classification scheme. In addition, some of the variables required for model parameterization were not available in the FAO database and had to be acquired from other sources. The methodology used to translate information from the FAO soils dataset and other sources into input parameters for hydrologic models will be presented, along with a comparison of model results and parameters using the FAO, SSURGO and STATSGO soils datasets.

Keywords: FAO; soil map; hydrologic modeling; GIS; watershed management.

Introduction

Water resources are a global concern as population grows and the need for water supplies expands. Watershed runoff and erosion models, such as the Kinematic Runoff and Erosion Model (KINEROS2) and the Soil and Water Assessment Tool (SWAT), are used to assess land-use and climate-change impacts on water yield and quality at multiple scales. When incorporated into a geographic information system (GIS) tool such as the Automatic Geospatial Watershed Assessment Tool (AGWA), model parameterization becomes much easier, and the use of complex hydrologic models for management and planning is facilitated.

AGWA was originally created for applications within the United States, and uses either the STATSGO or SSURGO soils data sets available only for the United States. As a result, analysis of basins that straddle the U.S. border, such as the San Pedro River basin in Southern Arizona, or are outside of the United States was not possible. In January 2003, version 3.6 of the Food and Agriculture Organization of the United Nations (FAO/UNESCO) Soil Map of the World was released. This digital soils data set includes

[1] USDA-ARS Southwest Watershed Research Center, 2000 E. Allen Rd., Tucson, AZ 85719. Respectively, llevick@tucson.ars.ag.gov, iburns@tucson.ars.ag.gov, sscott@tucson.ars.ag.gov, cunkrich@tucson.ars.ag.gov, dgoodrich@tucson.ars.ag.gov
[2] U.S. E.P.A. Landscape Ecology Branch, PO Box 93478. Las Vegas, NV 89193, semmens.darius@epa.gov
[3] University of Arizona, School of Natural Resources, Tucson, Arizona 85721, phil@nexus.srnr.arizona.edu

soil maps and properties for the entire world. The FAO soils data set is currently being added to AGWA to make it applicable worldwide and across country boundaries.

This paper describes the integration of FAO soils into AGWA, and compares the parameter set created by AGWA and the runoff results from KINEROS2 for the FAO, STATSGO and SSURGO soils data sets. While AGWA can also parameterize the SWAT model with FAO soils, SWAT results and analysis will be presented at a later time.

General Description of the FAO dataset

The FAO Soils dataset was originally published in 1974 as the Soil Map of the World (SMW). It was first digitized by Environmental Systems Resources Institute (ESRI) in 1984 in vector format, and later re-projected and converted to grid format by the Global Resources Information Database (GRID) project of the United Nations Environmental Programme.

At that time, it contained 26 “Major Soil Groupings” with 106 “soil-units”. In 1990, a new legend was published that introduced a third level of “soil subunits”. The International Union of Soil Sciences (IUSS) officially adopted this World Reference Base for Soil Resources (WRB) as their soil correlation system in 1998. Currently, there are 30 Reference Soil groups organized into 10 sets. Organic soils are grouped into set number one, and mineral soils are in the remaining nine sets (Driessen et al., 2001).

The Digital Soil Map of the World includes a total of nearly 5000 mapping units and more than 200 soil-units. The dataset contains digital map layers in ArcInfo GIS export format, and associated tables of soil properties for the entire world. The SMW was compiled from individual country soils data, which used a variety of local soils data, and is a generalization of those data. The map scale is 1:5,000,000. In comparison, the SSURGO maps are at scales ranging from 1:15,000 to 1:31,000. STATSGO maps, at a scale of 1:250,000, were created by generalizing the SSURGO maps (Bradley, 2003). The FAO maps are arranged into 10 major continental regions: Africa, Australasia, Central America, South America, Europe and West of the Ural, North America, Central and North East Asia, Near East, Far East, and South East Asia. Country boundaries have been updated as of 1994 at 1:3,000,000 scale (FAO/UNESCO, 2003).

Distributed hydrologic models that compute water balances, simulate climate, or estimate crop growth require information about the soil hydraulic properties, some of which is not available in the FAO soils database. Unavailable soil properties include available water-holding capacity (AWC) and organic carbon for SWAT, and rock fragments and top soil depth for both SWAT and KINEROS2. To address this need, Reynolds et al. (1999) estimated available water-holding capacity for FAO soils using continuous pedo-transfer functions within global pedon databases and linked these results to the SMW. While estimating AWC, they also determined and summarized rock fragments, organic carbon, and soil depth using additional sources such as Batjes (1997).

FAO soil maps are comprised of polygonal *mapping units*. Each mapping unit is associated with a unique number in the feature attribute table that is referenced to the FAO SU.ASC table, which contains the FAO soil identifier name and number, and the *soil-units* contained within the mapping unit (Figure 1). Specific properties associated with each soil-unit are stored in separate tables for each property. Reynolds et al. (1999) summarized many of these properties, making the information much easier to access and use in AGWA.

Methods

AGWA uses soils, land cover, and topographic (digital elevation model) data layers to obtain input parameters for its two component hydrologic models, KINEROS2 and SWAT. AGWA will first delineate the study watershed based on the user specified watershed outlet location, and then discretize, or subdivide, the watershed into model elements: subwatersheds (SWAT) or planes (KINEROS2), and channels. The subdivided watershed map is then intersected with the soil and land-cover data to extract the required model input parameters for each watershed element. A supplemental look-up table is necessary to derive parameters that are not included in the soil and land-cover datasets.

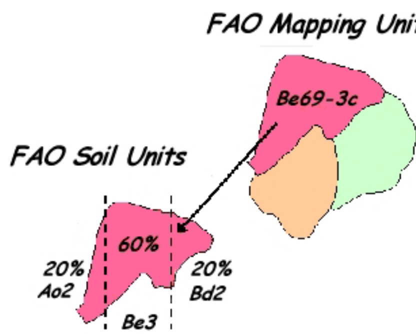

Figure 1. FAO Mapping Units and Soil-units

The structure of the FAO Soils database is fundamentally different from STATSGO or SSURGO. Therefore, adding FAO soils to AGWA required creating four new look-up tables that contained data from the FAO soils and Reynolds databases. When retrieving parameters for use as inputs to the models, parameters were obtained first from the FAO data if available, then from the Reynolds data, and finally from the AGWA table if there was no other source. The GIS data layers for the FAO soils were not altered other than to project them to a local projection (e.g. UTM Zone 12) for testing. The four tables are described below.

1. Component tables: FAO_World.dbf and FAO_Africa.dbf. The FAO soils database divides the soil mapping units into Africa and the rest of the world. The two tables created for AGWA are organized by soil mapping unit number and FAOSOIL name. They include the soil-units and percentages associated with each soil mapping unit. These tables were derived from the SU.ASC table, from the FAO Soils data CD (FAO/UNESCO, 2003).

2. Properties table: FAO_Properties.dbf. This table lists the soil-units and properties (% sand, % silt, % clay, bulk density, etc.) for the top and sub soils. FAO soil-units consist of a top-soil generally of 30 cm depth, and a sub-soil to 100cm depth. This table was derived from the SUSTXT.txt table, from the FAO Soils data CD (FAO/UNESCO, 2003). A texture description was needed for AGWA but was not present in the original FAO table, so we added a “texture” field, and populated it using the sand-silt-clay fractions and the USDA Soil Texture Triangle. Texture is used to obtain hydraulic parameters, such as Ks, from the look-up tables described below.

3. Summary Table: FAO_Summ.dbf. This table was obtained from the Reynolds, et al., (1999), summary table and is organized by soil-unit. It contains additional soil-units and soil properties not found in the original FAO database such as AWC, and rock fragments. It was not modified for use in AGWA other than converting it to dbf format and adding the same “texture” field described for the properties table above.

4. The kin_lut.dbf table from the AGWA database consists of soil textures (i.e. sand, loam, clay loam) and their associated soil properties such as saturated hydraulic conductivity (Ks), mean capillary drive (G), and porosity (por). AGWA uses soil-unit texture to obtain soil properties for model parameterization. We were able to give most FAO soil-units a texture class from the soil texture triangle for accessing the kin_lut.dbf. However, the FAO dataset contains some soil mapping units that do not

have any associated soil-units or soil properties. This table was modified to include these unique FAO mapping units: dunes/shifting sands, salt, rock, glaciers, water and No Data. These mapping units were treated as textures, and assigned parameter values based on the most similar soil texture in the table. For example, “salt” was given the same values as “gypsum”, while “rock” was given values from “unweathered bedrock”.

Occasionally a soil-unit is not found in the FAO_Properties.dbf table, so texture was determined based on the soil-unit name. FAO soil textures are “coarse”, “medium” or “fine”, and are designated with a “1” (coarse), “2” (medium), or “3” (fine) following the soil-unit symbol (i.e. AF 1). These textures are defined in the FAO documentation based on the soil textural classes from the USDA Soil texture triangle (FAO/UNESCO, 2003). The three soil textures were added to the kin_lut.dbf, and their soil properties were created from weighted averages of values in the kin_lut.dbf. For example, “coarse” soil, defined as “Sand, Loamy sand, and Sandy loam soils when clay is less than 18% and sand is greater than 65%” (FAO/UNESCO, 2003), used the weighted average of these three soil textures (see Figure 2).

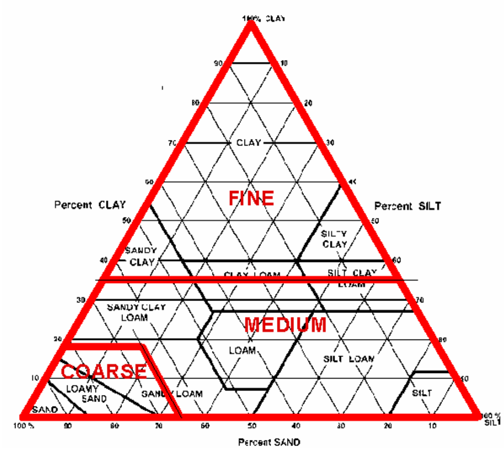

Figure 2. USDA Soil Texture Triangle with FAO soil textures outlined.

Model Parameterization

AGWA creates the input file for KINEROS2 using parameters obtained from the GIS coverages and look-up tables. Both the STATSGO and SSURGO soils datasets are organized by a unique soil-id for each soil polygon, which in turn consist of soil components, and layers or horizons. AGWA obtains some of the soil properties from the look-up tables provided with the digital maps. The rest are obtained from kin_lut.dbf based on texture. To get an average value for each of the KINEROS input parameters on each model element AGWA performs both depth and area weighting of texture properties (Burns et al., 2004)

FAO soils are described differently from STATSGO or SSURGO. Soils are organized by soil mapping units, which consist of up to eight soil-units. Soil properties are defined for the top soil and sub soil for each soil-unit. The top soil is assumed in most cases to be 30cm (11.8 inches) deep. KINEROS2 is concerned only with the top 9” of soil which has the most influence on runoff. Therefore, only the top-soil properties are used from FAO soils, and depth weighting is not required.

AGWA performs a series of calculations to derive the input parameters for the KINEROS2 model. The soil properties for each soil-unit are obtained by intersecting the soils map with the watershed map, retrieving the soil-id from each model element (planes), and then getting the soil properties from the look-up tables. The general logic and look-up tables used to do the soils weighting depend on which FAO soil is present. For example, the FAO mapping units *water*, *no data*, *rock*, *salt*, *glacier*, or *d/ss* have no associated soil-units, and there are no properties in the FAO database. Therefore, these mapping units were assigned parameters from the kin_lut.dbf table as described earlier. When the FAO soil is *d/ss* (dunes/shifting sands), all input parameters are obtained directly from the "sand" texture in the kin_lut.dbf, the default look-up table. If the FAO soil is *salt*, *glacier*, or *rock*, then all the input parameters are set to "0", except for the "pave" value (representing erosion pavement for erosion calculations, not used in runoff calculations), which is set to "1". If the FAO soil is *water* or *no data*, this mapping unit is not included in the calculations. All other mapping units have soil-units, so properties are obtained from the tables described earlier.

Comparison with STATSGO and SSURGO soils

Parameterization routines based on FAO soils were compared with those based on STATSGO and SSURGO by running all three on sub-watershed 2 at the USDA-ARS Walnut Gulch Experimental Watershed, Arizona (Figure 3), using the KINEROS2 model. Walnut Gulch is one of the most extensively monitored watersheds in the United States, with more than 100 instrumented locations (rain gages and flumes) over the 148 sq. km watershed.

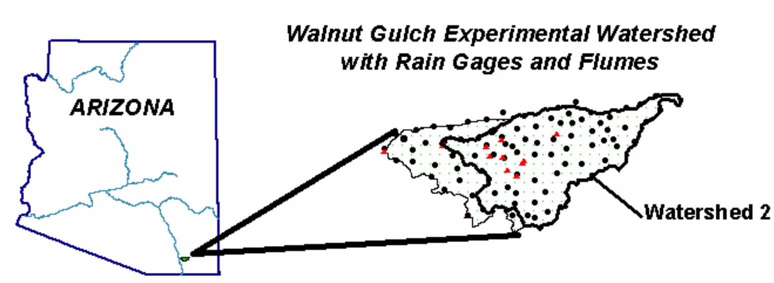

Figure 3. Walnut Gulch Experimental Watershed, Arizona

For these simulations, the land cover was set to a uniform "desertscrub" which comprises 55.5% of the watershed. This was based on research conducted by Bradley (2003), who compared the STATSGO and SSURGO soils databases for several watersheds at Walnut Gulch to determine the effects of soil data resolution on KINEROS2. By keeping land cover constant, the effects on runoff from land cover are reduced, allowing a better comparison of the soils data. It is important to note that the runoff results reported here would be different had the actual land cover been used, and would also allow comparison to the observed runoff data. In addition, the 28 observed precipitation events from Walnut Gulch used in the Bradley study were used for this study. A precipitation file produced from this network consists of rainfall depths for each rain gage that received precipitation during the rainfall event. These events ranged in size from 2.37 mm to 37.11 mm, with a mean of 16.0 mm and a standard deviation of 8.38. For comparison purposes, a 10-year 1-hour design storm for this region would produce a rainfall depth of 20.56 mm. All 28 events used in these simulations occurred during the summer monsoon season, from June through October, since these high-intensity convective thunderstorms produce the most runoff in this semi-arid desert/grassland environment.

Bradley (2003) simulated on 4 subwatersheds within Walnut Gulch, and found that "using the higher resolution SSURGO database increased the ability of KINEROS2 to predict runoff with respect to observed data." The present study used the largest of the subwatersheds (subwatershed 2, 108 km^2, Figure 3) from the Bradley (2003) study, the same 28 rainfall events, and the same uniform desertscrub land cover. Due to subsequent updates to AGWA and the KINEROS2 model since the Bradley study, the STATSGO and SSURGO simulations were repeated for this research. However, the trends and conclusions reached in that study will be noted here for comparison. It is important to mention that the model was not calibrated for these simulations.

Results

The bar graph below (Figure 4) shows the runoff results from each soil type. Only the events that produced runoff from at least two of the soil types, or were greater than 0.01 mm (10 of the 28 events) are shown. Runoff using STATSGO soils was generally higher than with SSURGO. This is consistent with the findings from Bradley (2003). In most cases, FAO soils produced less runoff than STATSGO soils (8 of the 10 events shown), although it produced more runoff than SSURGO soils about half the time (6 of the 10 events shown).

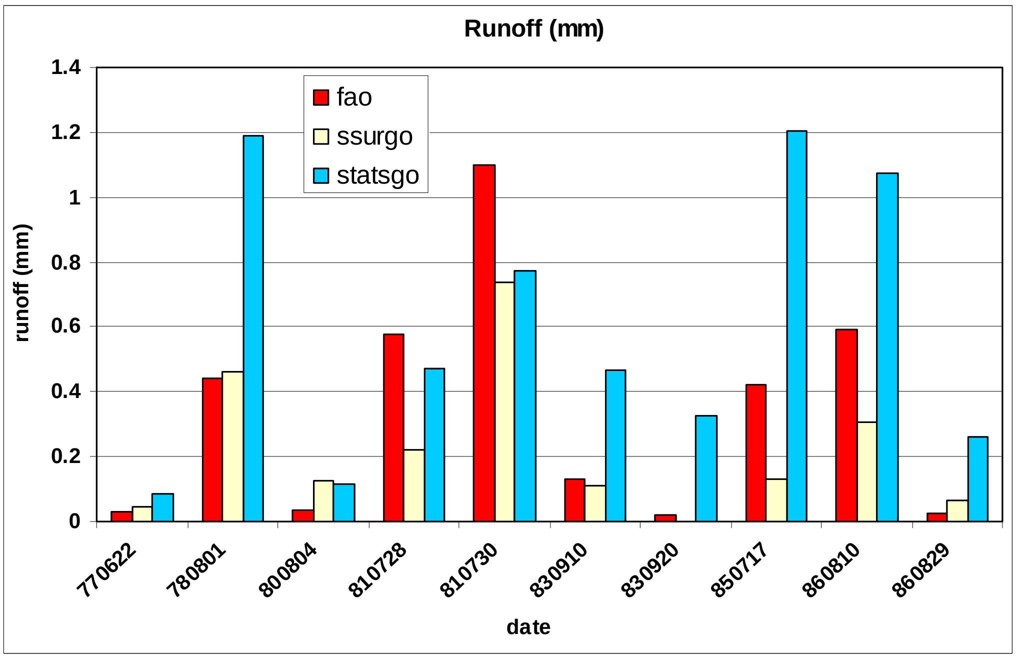

Figure 4. KINEROS2 simulated runoff (mm) at Walnut Gulch, Watershed 2, for FAO, STATSGO and SSURGO soils for selected rainfall events

Figure 5 is the regression analysis of runoff from FAO soils vs. STATSGO and SSURGO, using results from all 28 rainfall events. If model results were identical between the soil types, the slope of the regression line would be “1”. For FAO vs. STATSGO the slope was 1.2582, indicating that FAO generally produced less runoff than STATSGO. For FAO vs. SSURGO, the slope was 0.6125, indicating that FAO generally produced more runoff than SSURGO.

A comparison of the most sensitive soil properties (Ks and G) for the three soils databases is shown in Table 1. SSURGO soils had the highest mean values for Ks (10.52 mm/hr), while FAO had the lowest (5.84 mm/hr). The differences in the values between the three datasets may be attributed in part to the differences in spatial resolution and to the weighted averaging of the soil polygons for each sub-watershed element. SSURGO, with the highest spatial resolution, has 24 unique soil polygons in Watershed 2, while STATSGO has 3, and FAO has 2 (Figure 6). Each soil polygon will have unique soil properties. Table 1 shows that SSURGO has the largest standard deviation for Ks and G due to the greater number of unique soil polygons, while FAO has the smallest. The higher Ks values for SSURGO are due to this finer resolution which results in more soil polygons (and soil textures) being included in each model element. These polygons can include soils with high Ks values, such as sandy soils along stream courses, which do not show up in the lower resolution FAO and STATSGO datasets.

A similar trend is seen for mean capillary drive (G), where the means were 247.91 mm for FAO, 119.11 mm for STATSGO, and 164.63 mm for SSURGO. These values also reflect the difference in spatial resolution between the three datasets and the spatial averaging. The results for Ks and G are consistent with Bradley's (2003) findings when comparing STATSGO to SSURGO.

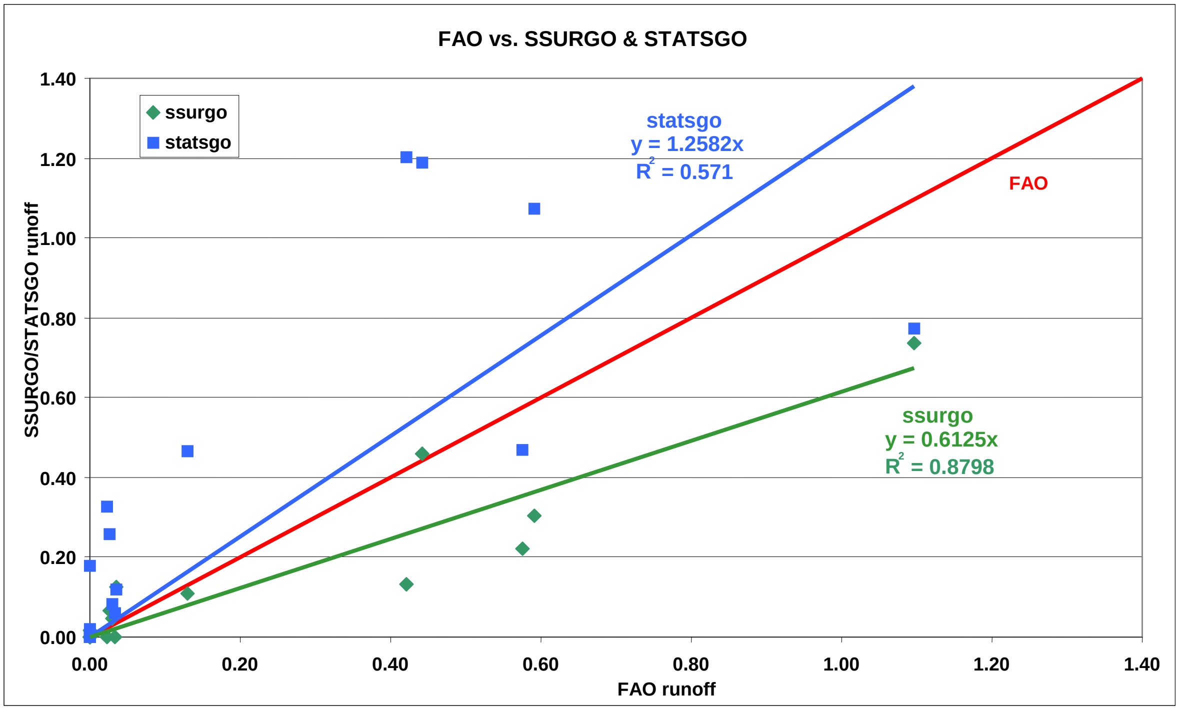

Figure 5. Runoff (mm) for FAO vs. STATSGO (a) and SSURGO (b) for 28 rainfall events

Table 1. Statistics of Ks (mm/hr) and G (mm) for Plane elements

	Ks (mm/hr)			g (mm)		
	FAO	STATSGO	SSURGO	FAO	STATSGO	SSURGO
Mean	5.84	8.23	10.52	247.91	119.11	164.63
Min	5.64	5.34	2.69	209.2	114.97	81.73
Max	7.66	8.44	105.52	252.2	195.03	262.25
Std.Dev	0.567	0.627	11.606	12.078	13.074	36.986

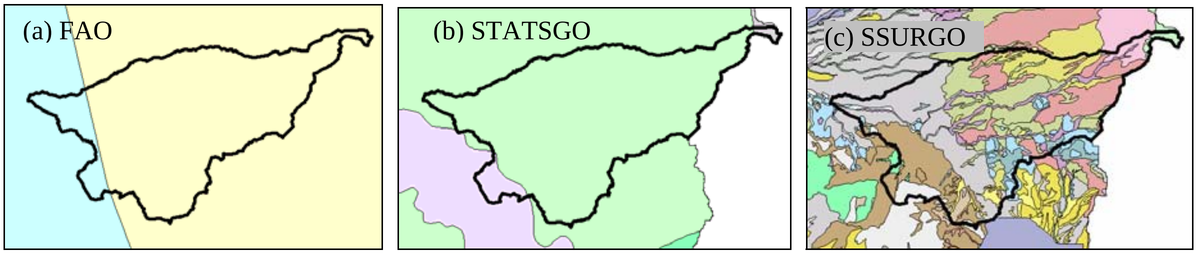

Figure 6. Watershed 2 outline with (a) FAO, (b) STATSGO and (c) SSURGO soils

Conclusions

The differences in runoff and soil properties can be attributed to the difference in the spatial resolution of the data sets. The SSURGO soils map has the finest resolution, while the FAO soils map has the coarsest. While STATSGO is a generalization of the SSURGO soil polygons, Bradley (2003) determined that the hydraulic parameters are not a spatial averaging of the SSURGO data. The FAO soils database was compiled from individual country soils data, which used a variety of local soils data, and is a generalization of those data. The results from this analysis indicate that the current integration of FAO soils into AGWA is adequate for hydrologic modeling, producing results comparable to the STATSGO and SSURGO soils data. Future analyses of the integration of FAO soils into AGWA will include comparison of simulated runoff to observed runoff using the actual land cover data.

References

Batjes, N.H., A world dataset of derived soil properties by FAO-UNESCO soil unit for global modeling, Soil Use and Manage., 13:9-16, 1997.

Bradley, C.M., Effects of soil data resolution on modeling results using a physically based rainfall-runoff model. M.S. Thesis, University of Arizona, Tucson, Arizona, 2003.

Burns I.S., S.N. Scott, L.R. Levick, M. Hernandez, D.C. Goodrich, S.N. Miller, D.J. Semmens, W.G. Kepner. Web site. Automated Geospatial Watershed Assessment (AGWA) - A GIS-Based Hydrologic Modeling Tool: Documentation and User Manual, Version 1.4, 2004. http://www.tucson.ars.ag.gov/agwa/

FAO/UNESCO, Digital Soil Map of the World and Derived Soil Properties, CD-ROM, Version 3.6. Information Division, FAO, Viale delle Terme di Caracalla, 00100 Rome, Italy, 2003

Driessen, P., Deckers, J., Spaargaren, O., Nachtergaele, F., editors. Lecture Notes on the Major Soils of the World. FAO, Publishing and Multimedia Service, Information Division, Viale delle Terme di Caracalla, 00100 Rome, Italy, 2001. http://www.fao.org/DOCREP/003/Y1899E/Y1899E00.HTM

Reynolds, C.A., T. J. Jackson, and W.J. Rawls. Web site. Estimated Available Water Content from the FAO Soil Map of the World, Global Soil Profile Databases, and Pedo-transfer Functions, 1999. http://www.ngdc.noaa.gov/seg/cdroms/reynolds/reynolds/reynolds.htm